LEVEL II

# SENSORY INFORMATION PROCESSING AND SYMBOLIC COMPUTATION

1 JULY 1973 THROUGH 31 DECEMBER 1973

Semi-Annual Technical Report



Contractor: University of Utah

Contract Number: DAHC15-73-C-0363

Effective Date of Contract: 1 July 1973

Expiration Date of Contract: 30 June 1975

Amount of Contract: $2,275,000.00

Project Code: 3D30

Principal Investigator: Dr. Thomas G. Stockham, Jr.
Telephone: (801) 581-8224

Contracting Officer: Mr. Eugene W. Stubbs
DSSW

Approved for public release;
distribution unlimited.



"Original contains color plates: All DTIC reproductions will be in black and white"



Sponsored by
Defense Advanced Research Projects Agency
ARPA Order Number 2477

362900

PUBLICATION REVIEW

This technical report has been reviewed and is approved.

Project Engineer

TABLE OF CONTENTS

## LIST OF FIGURES

## REPORT SUMMARY

This Semi-Annual Technical Report covers the research performed between 1 July 1973 and 31 December 1973 in the areas of Sensory Information Processing, Symbolic Computation and Graphics.

The Sensory Information Processing research develops and applies innovative digital signal processing techniques--processing two-dimensional (video) signals and one-dimensional (audio) signals. The efforts currently emphasized and reported herein include the study of automatic deblurring of images, linear predictive coding of speech and the refinement and application of mathematical models of human vision and hearing. All of these efforts employ a general purpose digital computer, although some (such as the linear predictive coding project) are directed toward the production of techniques amenable to implementation with special purpose digital hardware.

The Symbolic Computation research is concerned with the development of automatic aids for the algebraic and symbolic computations inherent in scientific research. In the course of this research, the REDUCE system has been developed and is being improved and distributed. Accordingly, this effort deals directly with questions of transferability of software between machines and the automatic generation of programs.

The Graphics effort is charged with the development of a functioning hardware and software system for the generation of shaded pictures from mathematical descriptions.. During the reporting period, that system was designed at the block level and procurement of its components was initiated. However, certain specific design choices have been left open and research into object modeling techniques and methods of interaction is continuing.

# SECTION A

# SENSORY INFORMATION PROCESSING

This research embraces a broad spectrum of topics ranging from problems associated with digitizing and undigitizing image and audio signals to speculative inquiry requiring sophisticated and innovative computer processing methods based upon psychophysics of human vision and hearing. Emphasis has been placed on obtaining very high performance quality, demonstrating the feasibility of processes which represent technological breakthroughs, the evolving of the computer as a basic tool for processing information in the form of signals and emphasizing the importance of human observer characteristics in producing successful processes. The following work represents our continuing efforts in these directions.

## Image Deblurring

Our research into the problem of deblurring images which have been corrupted by unknown amounts and types of blurs has paid off very well recently. The work of Cole[1] has been extended to include two important practical cases involving phase shifts of the spatial frequency components of an image. These cases are uniform motion blur and defocus blur. In addition, long exposure atmospheric turbulence blur has been considered. All three cases have been tested in the laboratory for actual "in-camera" blurs. See examples in Figures 1, 2, 3, and 4.

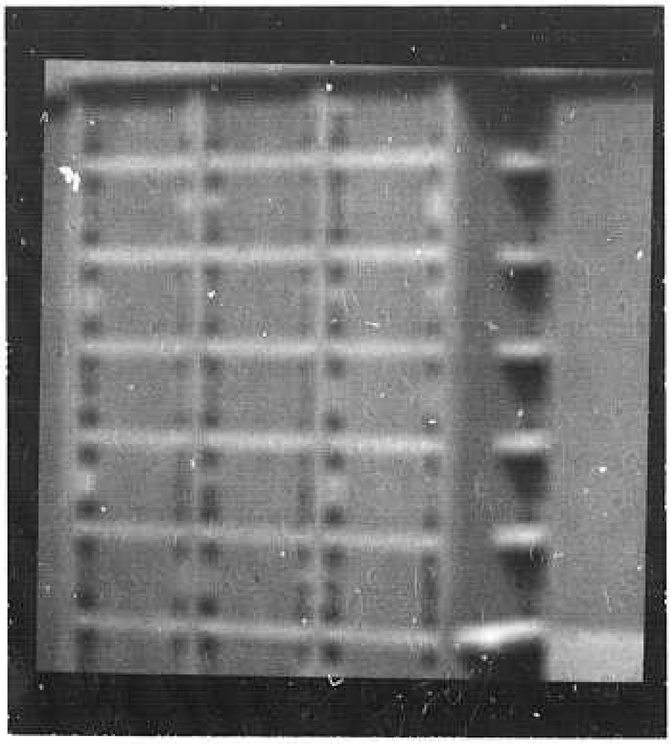


Figure 1

Original Out-of-Focus Photograph

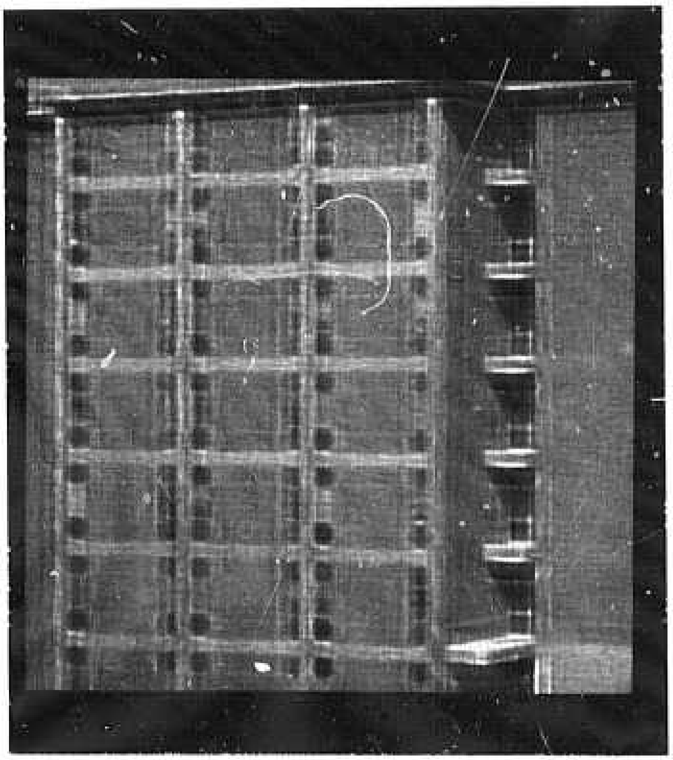


Figure 2

Enhanced Logarithmic Restoration


Figure 3

Original Motion-Blurred Photograph


Figure 4

Logarithmic Restoration

Although applicable to intensity images, the method as demonstrated here, has been applied to log-intensities. This step insures that positive definite deblurred intensities are finally obtained and allows for combined homomorphic enhancement of the restored image.

The method used deduces the phase required for the restoration by carefully noting the apparent zero crossings of the power spectrum of the blurred image. While Cole used the average log-magnitude-spectrum to estimate the amplitude distortions of his unknown blurs, the computation times required for our deblurring scheme were reduced by substituting the log-power-spectrum estimate instead. The results were found to be subjectively equivalent in practice. This fact is partially predicted by the theoretical studies reported under the section of this report entitled "Homomorphic vs. Power Spectrum Estimation of Amplitude Distortion Functions." The time saving comes from the use of the logarithm function only once after a power spectrum has been estimated instead of many times as when averaging log-magnitude-spectra.

One phenomenon which seems to plague all image deblurring methods when working with high signal to noise ratios, is that of a ringing or ghost image phenomenon which surrounds high contrast edges. Methods for avoiding such ringing for objects which are known to be surrounded by a featureless field, such as the sky, has been demonstrated.

During the next reporting period, research will center upon the problem of finding the phase by methods that are more reliable and automatic than the zero crossing detection method already in use. It is also hoped that combined motion and defocus blur will become part of our automatic deblurring repertory. It is also very likely that report writing will commence on this project.

A considerable interest in the Utah methods for automatic recovery from unknown blurs has been demonstrated by the technical community including both government agencies and their contracting organizations.

## Eye Model for Brightness

The homomorphic model for predicting brightness perception in the human visual system, as developed by Stockham [2], is a major result of sensory information processing research at the University of Utah. Subsequent work by Colas-Baudelaire [3] refined this model but also pointed out some weaknesses in it. It is the purpose of this research to modify the model in order to prevent excessive sharpening of object boundaries which characterizes the homomorphic model and detracts from its usefulness in the most precise and demanding situations.

The homomorphic model consists of a point by point logarithm followed by a two-dimensional linear filter. The linear filter is then followed by a saturation element as shown in Figure 5. The above mentioned difficulty with this model is traceable to the step response of the linear filter. This step response correctly predicts brightnesses that are observed subjectively for smooth patterns.

The problems obtained at object boundaries are illustrated in Figures 6a, 6b and 6c. In Figure 6a, a step function serves as a simplified representation of an object boundary. Figure 6b and 6c show the subjective brightness predicted by the homomorphic model and the subjective brightness reported by typical observers respectively. During the past six months, heavy attention has been given towards formulating a variety of alternate models which might have an opportunity for overcoming this effect while retaining the desirable properties already obtained. No completely satisfactory alternative has been discovered. However, a model with considerable promise has been studied with the hope that some extensive practical experience with its effect upon real images would suggest a more complete and satisfactory theory.

This tentative model has the following properties. First, it divides the image into a rough and smooth part by differentiation followed by a resistive filtering operation [4]. Next, a linear filtering step is applied to the smooth part and finally the two components are added and integrated to obtain an output. This model is shown in Figure 7. Note that it reduces to the homomorphic model when sharp contours, such as shown on Figure 6a, are not present because the smooth part becomes the same as the differentiated log-intensity.

During the next reporting period this new model will be applied to real images and its performance evaluated.

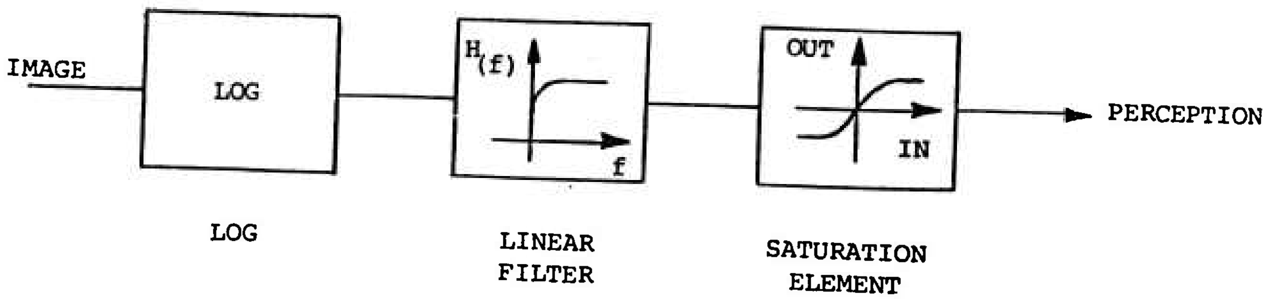



Figure 5

Homomorphic Model of Human Brightness Perception

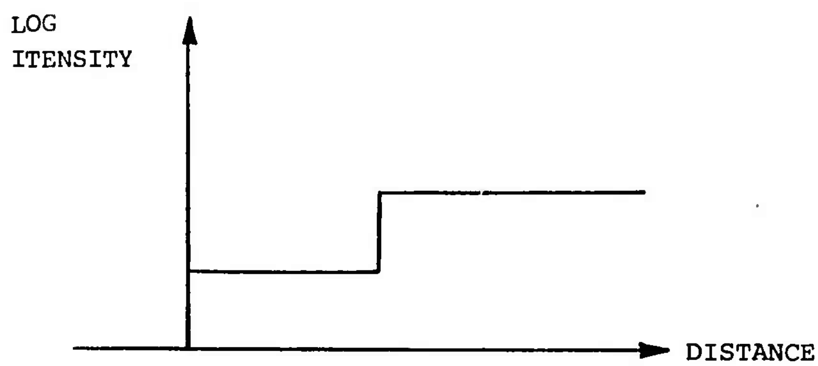



Figure 6a

Test Image

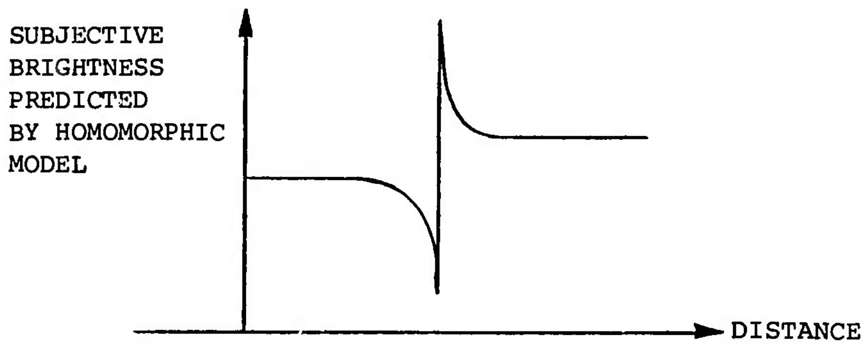



Figure 6b

Model Response

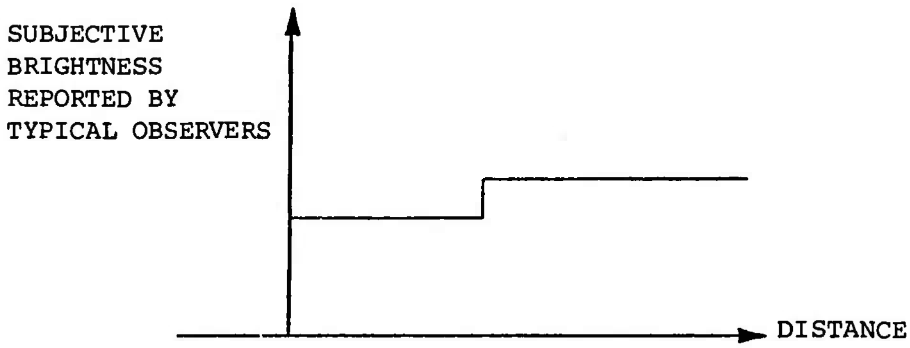



Figure 6c

Visual System Response

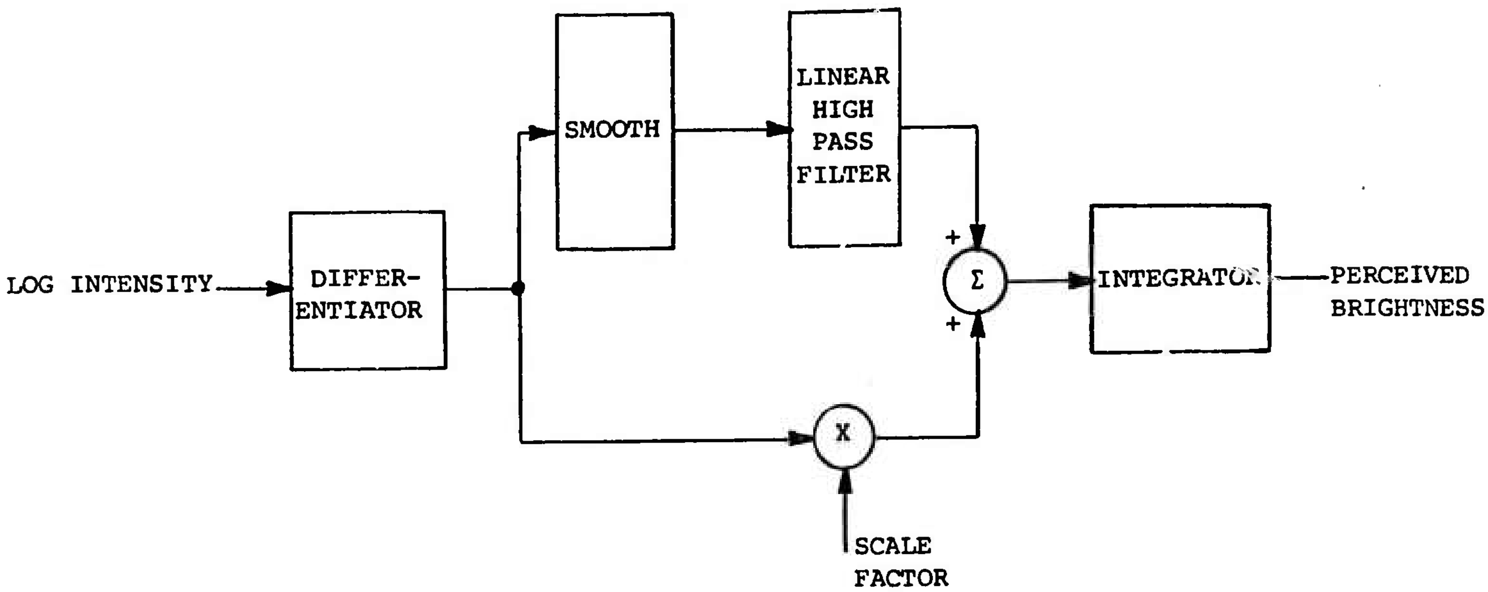



Figure 7

Modified Visual Model

## Color Correction of Digital Images

The problem of displaying high quality color images produced by a digital processor is a difficult one. A color television monitor refreshed in real-time from a digital frame buffer memory can provide excellent colorimetric results with quick look capability and medium resolution. Permanent recording and high resolution presently require the use of color photography, however.

Unfortunately, the colorimetric fidelity of color photography is low due to the imperfect nature of the colored dyes which must be used. Lithographers have long known of this problem and have developed methods for compensating the dye imperfections. These methods are simple and effective, but have been generally overlooked by digital image processors.

The goal of this research has been to integrate the basic principals used in lithography into the software of our high-quality digital image recording system without sacrificing the quick turnaround and tight quality control already obtained for black and white. Once developed, these techniques would be applicable not only to natural scenes, but also to other types of digital data intended for display on color film.

Some preliminary results are illustrated in Figures 8 and 9. Although the printing used to make this report has not been proofed at the time of this writing and therefore may partially compromise the demonstration, the color recordings themselves show a marked improvement in color quality in the processed version. The colors have become more saturated and closer to their proper hues. This effect is most noticeable in the flowers, the scarf, and the dress color. Direct comparison of the processed version with the slide from which the digital data was originally scanned is very gratifying especially since the color recordings were made on Polaroid R material.

Work is continuing in order to make these methods usable on a routine basis in our image processing laboratory. It is hoped that experience will show that these methods can enjoy a much wider use in the image processing community.

Figure 8

Color Display Without Compensation

Figure 9

Color Display With Compensation

Homomorphic vs. Power Spectrum Estimation of Amplitude Distortion Functions

As discussed earlier in the section entitled "Image Deblurring," both average log-magnitude- spectra (i.e., homomorphic) and log-power-spectrum estimates have been used to estimate the amplitude distortion of unknown image blurs. While our experience so far shows that the two methods produce subjectively equivalent results when applied to the removal of unknown image blurs, such is not the case when applied to the removal of unknown resonances in sound. The question naturally arises as to why this difference should exist and which method is superior.

Our research has revealed that if one assumes that images and sound are stationary, gaussian, and white signals (a situation far from reality), that the homomorphic estimator converges slightly less rapidly (has about 50% greater variance) than the power spectrum estimator. However, the expected values of the two estimators differ only by Euler's constant (0.57721...) the power spectrum estimator being the greater. When used to determine amplitude distortions, the bias in the expected value cancels out leaving the homomorphic method equivalent to the power spectrum method except for a slightly larger uncertainty.

Theory and extensive experimentation carried out during this reporting period reveals that though images and sound are far from gaussian and white signals, the above results still hold approximately but with very small error, provided the stationary assumption is met.

If the signals in question are non-stationary (in this context images and sound must be considered non-stationary), experimental results indicate that the two estimation methods yield substantially different amplitude distortion curves. Using the power spectrum estimate to compute a deresonating system for old Caruso recordings yields an amplitude restoration frequency response that gives far more natural sounding results than the originally used homomorphic method. The frequency response also corresponds more closely to the amplitude distortions predicted from the available data on the frequency responses of old acoustic recording equipment.

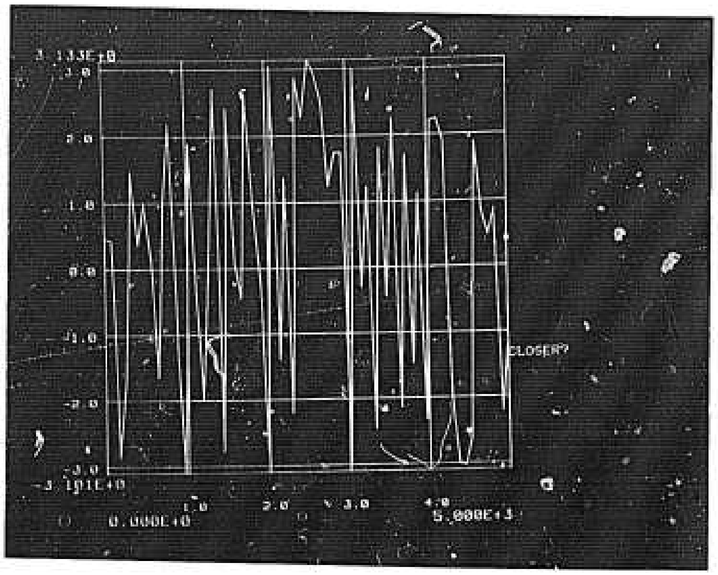
It remains to be shown why the non-stationary character of sound leads to these results and to assess the impact of these measured results on the image deblurring situation.

## Removing the Buzzy Quality from Synthetic Speech

With the recent increase in interest in the quality of synthetic speech that has been brought about by the problem of secure voice communications, we have begun to study the possible causes of the buzzy quality that all synthetic speech seems to have and which is considered by many to be a very objectionable factor. During the previous six month period, we have investigated the notion that the undesirably buzzy quality is caused by the instantaneous peak factor of the synthetic speech wave as compared with natural speech wave and that all-pass filters with random phase vs. frequency responses would tend to relieve the situation. In order to accomplish this, we synthesized several digital all-pass filters with various amounts of random phase and used them to process a small set of synthetic speech samples which were produced by an LPC vocoder and had unusually high quality except for their obvious buzzyness when auditioned through headphones. After all-pass filtering with these filters, the buzzy quality of the speech became less noticeable. However, the processed speech seemed to suffer a degradation, the extent of which depended upon the phase vs. frequency response chosen. Figures 10 and 11 show the characteristics of one of these all-pass filters.

These results, while disappointing, have focused our attention on other possible sources of the buzzy quality and upon the continued need for better understanding of how the human ear is sensitive to phase distortions.

PHASE (RADIANS)

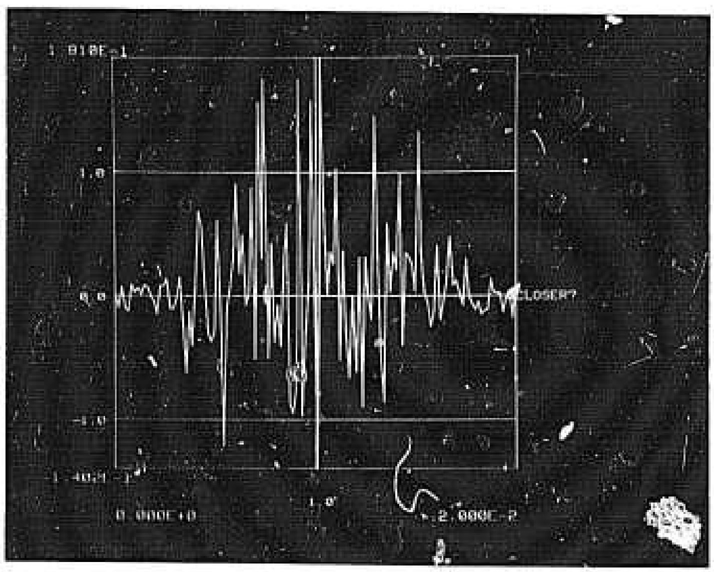



FREQ (HZ)

Figure 10

A Random Phase Configuration

IMPULSE RESPONSE




TIME (SEC)

Figure 11

The Impulse Response of an All-Pass
Random Phase Filter

## Linear Predictive Coding

Because of their importance in a real-time secure speech communications system, the speech coding research at Utah has concentrated upon the issues of quality and naturalness. The goal of this research during this period was to continue to investigate methods for improving speech quality, especially by means of coefficient smoothing, while meeting the constraints of (1) narrow channel bandwidth transmission rate, that is, 3600 bits/sec or less; and (2) low computation requirement suitable for real time implementation by special purpose hardware.

Computer programs have been written for simulating a complete linear predictive coding voice analysis-synthesis system. Included are algorithms for estimating the vocal tract parameters based upon the all-pole model using either the covariance or the auto correlation method, algorithms for estimating pitch and voicing using a modified auto correlation method, and algorithms for simulating channel bandwidths from 9600 bits/sec to 1200 bits/sec.

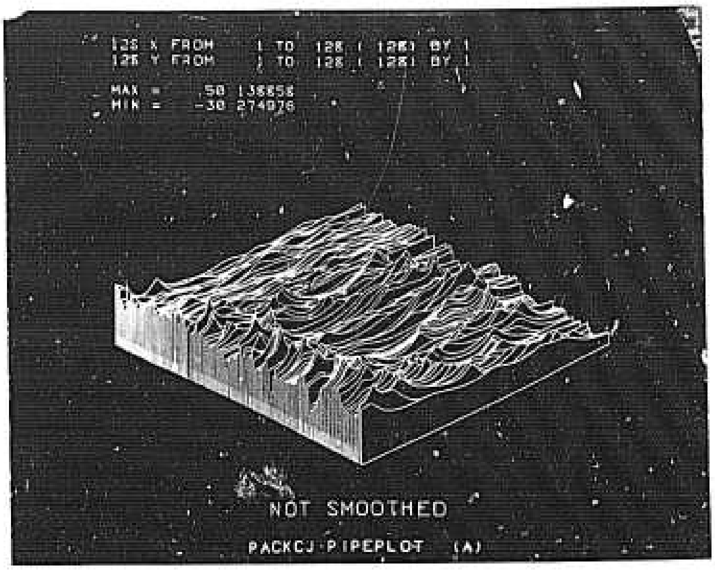
Using these programs, research was continued to determine the degree of improvement in speech quality obtained by performing additional smoothing of the time histories of the channel parameters prior to synthesis. It has been shown that smoothing the vocal tract parameters can be achieved by estimating either the predictor or reflection coefficients using an a priori least squares estimator. The degree of improvement increases as the channel bandwidth decreases. It was determined that smoothing improved synthetic speech intelligibility and naturalness for channel bandwidths less than 3600 bits per second.

The technical development of smoothing predictor coefficients is presented in Steven F. Boll's technical report [5]. An example of the effect of smoothing applied to predictor coefficients is given in Figure 12. In this figure, log magnitude spectra obtained with and without smoothing is displayed.

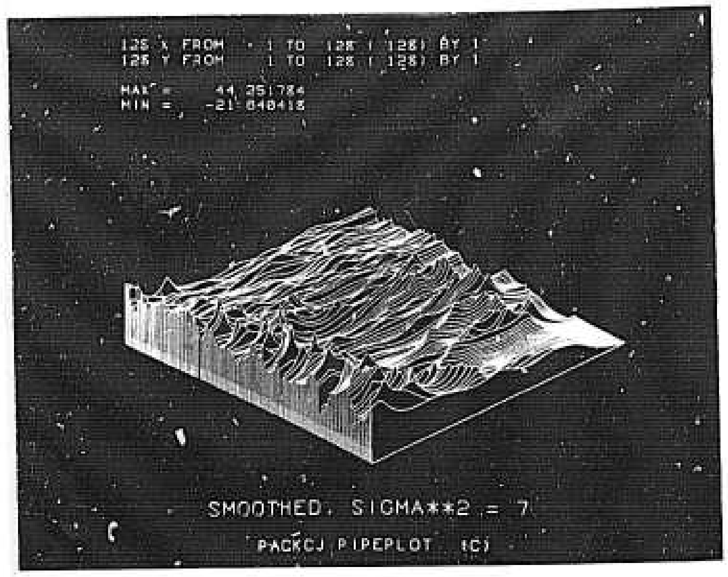
When estimating reflection coefficients using a priori least squares, the resulting coefficients also exhibit smoother time histories than those obtained from standard methods. Figure 13 shows the time histories of the first PARCOR coefficient with and without smoothing from analyzing the connected phonemes: /ei/, /i/, /I/, /ou/, and /u/. It can be shown that smoothing results from the fact that each coefficient is passed through a time varying, single pole low pass filter. The pole location for each filter, as determined by the a priori least square operator, varies as a function of the type of speech being analyzed, having a value near one for sustained vowels and diphonges and a value near zero for stops and transitions. Thus the estimator will only smooth the coefficients when the vocal tract shape is relatively constant or slowly varying. The result is a more natural synthesis speech quality but without loss of intelligibility, since essentially no smoothing is applied during transition regions.

Normally, the implementation of automatic smoothing using a priori least squares requires the need for matrix difference equations. However, by incorporating a Gram-Schmidt diagonalization method developed by Mitsui, the matrix difference equations are diagonalized to a set of uncoupled scalar difference equations. With this simplification, the total amount of arithmetic computation is sufficiently reduced to admit to real-time implementation.

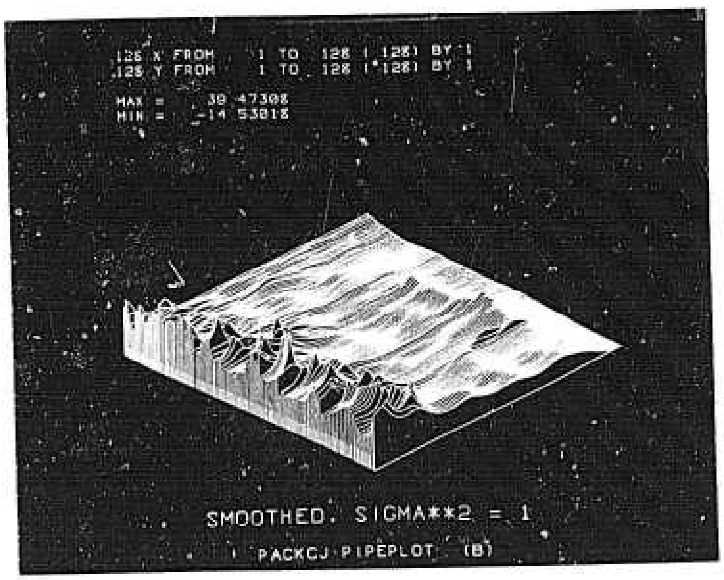
Thus, the results of this study show that the a priori least squares estimator, normally only associated with modern control theory applications, can be used to improve linear predictive coding synthesis speed quality without unreasonable increases in computation.




(a)




(b)




(c)

Figure 12

(a) Unsmoothed Log Spectra of the Sentence "The pipe began to rust." (b) and (c) Smoothed log spectra.

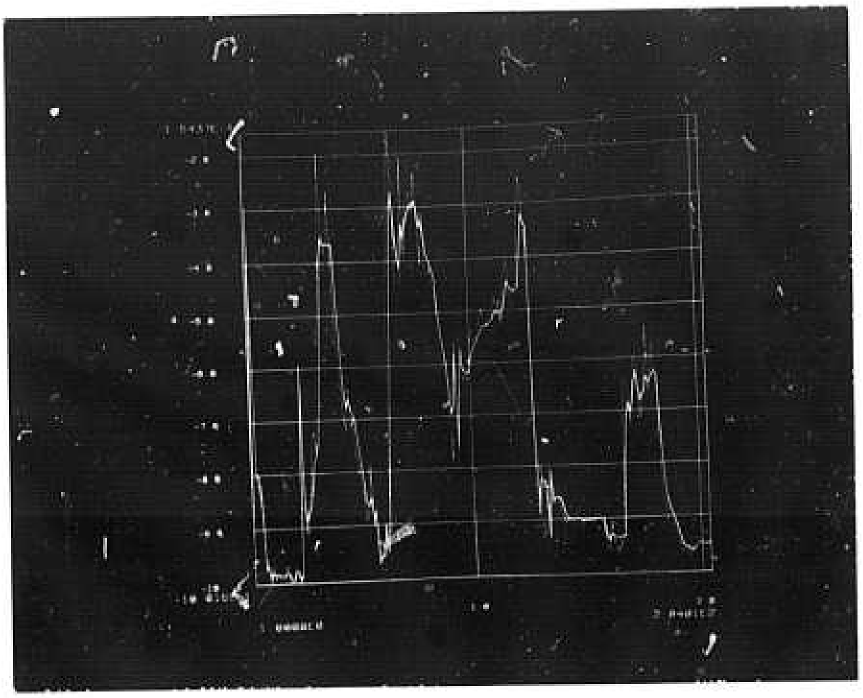


Figure 13

Smoothed and Unsmoothed $k_1$ Reflection Coefficient

# SECTION B

# GRAPHICS

The major part of the work carried out during the reporting period has involved a continuation of previous investigations. A significant event was the decision not to pursue negotiations towards subcontracting a major portion of our development work. In spite of this, work is progressing towards a system of reduced, but still significant, capabilities.

## Shaded Picture System

The PDP-11/45 computer, which forms the kernel of the developing Shaded Picture System, has been installed, and purchasing procedures for other parts of the system have been initiated. Modifications have been made to the DOS Monitor to provide access to more than the standard 28K memory.

## Display of Curved Surfaces

Techniques for smoothing the contours of objects represented with planar faces is continuing. The approach has progressed from attempting to smooth edges in the picture space, which can give rise to inconsistencies if the view point is moved. The present approach involves replacing the planar faces which make up the contour edges with bicubic surfaces, and then rendering the curved surfaces, which will not change as the viewpoint is moved. This has meant facing up to the problem of rendering curved surfaces directly, which we have been avoiding up to now, but which we feel is essential if significant advances are to be made.

## Modeling of the Human Face

The work on parameterization of the human face has yielded a 25 parameter model. This model facilitates the manipulation of facial expressions, including lip movements necessary for realistic speech synchronization. The present parameters control such things as eye opening, pupil dilation, mouth shape and eyebrow position. The eyes automatically track a specified point in space.

## Dynamic Line Drawing System

The reconstruction of the line drawing system is nearing completion. The new clipping-divider is almost complete and will enter testing early in 1974. The head mounted display has been reconditioned and is working, as is the three wire three dimensional input wand.

## Curved Surface Mathematics

A 3-D computer aided geometric design system is being developed that is based on surface representations due to Barnhill and Gregory [6]. The purpose of this work is to obtain better representations of curved surfaces than are currently available. Current methods are based on subdividing a region into squares and interpolating (or approximating) over each square with Coons patches or tensor product methods. Line drawing displays of a smooth rational interpolating surface patch over a triangle have been obtained.

Raster Scan Display of Line Drawings

Some preliminary investigations into producing line drawings on comparatively low resolution raster scan displays have been started, using our precision display equipment. This work is of interest since we would like to use our proposed digital Frame Buffer for the display of line drawings as well as half tone images. Initial results indicate that the technique of assuming lines have thickness and finding what proportion of each raster point is covered in order to establish intensity does not give optimal results, a prominent 'beading' effect being evident in a number of cases. This effect is being further investigated.

Application of Shaded Picture Technology to Vibration Studies

Associated DOD work on analizing and displaying wing flutter in supersonic delta wing aircraft has produced a convincing demonstration of the worth of our shaded picture generating capability. A final movie with sound commentary is being produced.

# SECTION C

# SYMBOLIC COMPUTATION

This group's research is directed toward the development of techniques for the solution of a wide range of symbolic and algebraic problems and the work during the reporting period may be divided into three main areas as described below.

## System Development

During the reporting period several major improvements in the REDUCE system have been initiated. In particular a complete re-writing of the substitution programs has been completed and tests indicate that significant improvements in the running times of many programs will result from these changes. For example a plasma physics program in use at the Max Planck Institute in Munich, Germany which took 500 seconds on an IBM 360-91 using the version of REDUCE released 18 months ago now takes 320 seconds on a PDP-10 using the latest version. This represents a factor of approximately 8 improvement in running time when machine differences are taken into account. Our work on improvements in the release of REDUCE for IBM computers continues as most exported copies of REDUCE are for these machines. Segmentation techniques are being implemented in order to run the program in a smaller core partition. This will increase its use at many installations because of improved turn-around time.

Language Standardization

We are continuing our efforts at standardizing the REDUCE language syntax and semantics. This work will make the program easier to use at remote installations on the ARPANET and elsewhere. We are currently studying the mode facilities of ALGOL 68, as a variation of these appears to offer the best possibility for semantic improvements in REDUCE. Some useful results in this area should be possible within the next six months or so.

Applications Research

We are continuing to study practical applications of algebraic techniques to problems in physics and engineering. However, more recently the theorem proving group at USC Information Sciences Institute has expressed interest in using REDUCE in their work. We are actively helping this group at the present time via the ARPANET. Another cooperative project was begun during this period with Dr. David Barton in Cambridge, England, via the London TIP. With the extension of the ARPANET to more and more scientists we anticipate even greater collaboration of this type in the future.

BIBLIOGRAPHY

[1] Cole, Edwin Randolph. "The Removal of Unknown Image Blurs by Homomorphic Filtering," Technical Report UTEC-CSc-74-029, University of Utah, Computer Science, June 1973.

[2] Stockham, T.G., Jr. "Image Processing in the Context of a Visual Model," Proceedings of the IEEE, Vol. 60, No. 7, July 1972.

[3] Colas-Baudelaire, Patrick. "Digital Picture Processing and Psychophysics: A Study of Brightness Perception," Technical Report UTEC-CSc-74-025, University of Utah, Computer Science, March 1973.

[4] Tukey, J.W. "Nonlinear (Nonsuperposable) Methods for Smoothing Data," Presented at the Workshop on Future Directions in Signal Processing, Annapolis, Maryland, July 1973.

[5] Boll, Steven F. "A Priori Digital Speech Analysis," Technical Report UTEC-CSc-73-123, University of Utah, Computer Science, March 1973.

[6] Barnhill, R.E. and J.A. Gregory. "Blending Function Interpolation to Boundary Data on Triangles," TR/14, Brunel University, Uxbridge, England, 1972.